



PCT

## RAW SEQUENCE LISTING

DATE: 09/21/2004

PATENT APPLICATION: US/10/507,106

TIME: 09:27:01

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

3 <110> APPLICANT: Japan Science and Technology Agency  
 5 <120> TITLE OF INVENTION: Gene participating in the synthesis of brassinosteroid  
 7 <130> FILE REFERENCE: 26352U (PS03-311PCT)  
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/507,106  
 C--> 9 <141> CURRENT FILING DATE: 2004-09-10  
 9 <160> NUMBER OF SEQ ID NOS: 6  
 11 <170> SOFTWARE: PatentIn version 3.1  
 13 <210> SEQ ID NO: 1  
 14 <211> LENGTH: 1473  
 15 <212> TYPE: DNA  
 16 <213> ORGANISM: Arabidopsis thaliana  
 18 <400> SEQUENCE: 1  
 19 atggacactt cttcttcact tttgttcttc tccttcttct tctttatcat catcgatcat 60  
 21 ttcaacaaga tcaacggtct cagatcatcc ccagcttcaa agaaaaaact taatgatcat 120  
 23 catgttacat cccagagtca cggaccaaag tttccacacg gaagcttggg atggcccgtc 180  
 25 atcggtgaaa ccatcgagtt cgtctcttct gcttactcag accgtcctga gagtttcatg 240  
 27 gacaagcgtc gtctcatgta tgggagagtg tttaagtcgc atatttttgg aacggcgacg 300  
 29 atcgtgtcga cggatgctga agtgaacaga gccgttttac agagcgactc gacagctttc 360  
 31 gtgccgtttt acccaaaaac ggtaaggagg ctaatgggaa aatcgtcgat acttcttatt 420  
 33 aacgggagtt tacatagacg gttccatgga ttagtcggtt ctttcttaaa gtcgccactt 480  
 35 ctcaaagctc aaatcgttag agacatgcac aagtttttgt cggaatccat ggatctatgg 540  
 37 tccgaggacc aacctgtgct cctccaagac gtctccaaga ctggtgcatt caaagtactt 600  
 39 gccaaaggcat tgataagtgt agagaaaagga gaagatttag aagagctaaa gagagagttt 660  
 41 gaaaatttca tatcaggact catgtcatta ccaattaact tccctggaac gcaactccat 720  
 43 agatctctcc aagctaagaa gaatatggtg aagcaagttg aaagaatcat agaaggcaaa 780  
 45 attaggaaaa caaagaacaa ggaggaagat gatgttattg caaaggatgt tgtggatgtg 840  
 47 ttgcttaagg actcaagtga acattttaact cacaatttga ttgctaacaa tatgatcgac 900  
 49 atgatgatcc ctggccacga ttctgtccct gtccctcatta cccttgccgt caaattcctc 960  
 51 tctgattctc ctgctgccct caatctccta acgaaaaaca tgaagctgaa aagtttgaag 1020  
 53 gaattgacag gagagccact atattggaat gactacttgt cgttaccttt aacacaaaag 1080  
 55 gtgattacag agacactgag aatgggaaat gttataattg gagtgatgag aaaggcgatg 1140  
 57 aaagatgttg aaataaaaagg atatgtgata ccaaaaaggat ggtgtttctt ggcctatctc 1200  
 59 agatcagttc atcttgatga agcttattat gagtctccgt acaaatttaa tccctggaga 1260  
 61 tggcaagaaa gggacatgaa cacgagtagt ttcagtcctt ttggaggtgg tcagagattg 1320  
 63 tgccctgggtc tcgatttggc tcgtcttgaa acttcagttt ttcttcacca tcttgtcact 1380  
 65 cgcttcagat ggatagcaga agaagacaca atcataaaact tcccaacggt gcatatgaag 1440  
 67 aacaaattac ccatttggat caaaagaata taa 1473  
 70 <210> SEQ ID NO: 2  
 71 <211> LENGTH: 490  
 72 <212> TYPE: PRT  
 73 <213> ORGANISM: Arabidopsis thaliana  
 75 <400> SEQUENCE: 2  
 77 Met Asp Thr Ser Ser Ser Leu Leu Phe Phe Ser Phe Phe Phe Phe Ile



## RAW SEQUENCE LISTING

DATE: 09/21/2004

PATENT APPLICATION: US/10/507,106

TIME: 09:27:01

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

```

78 1          5          10          15
81 Ile Ile Val Ile Phe Asn Lys Ile Asn Gly Leu Arg Ser Ser Pro Ala
82          20          25          30
85 Ser Lys Lys Lys Leu Asn Asp His His Val Thr Ser Gln Ser His Gly
86          35          40          45
89 Pro Lys Phe Pro His Gly Ser Leu Gly Trp Pro Val Ile Gly Glu Thr
90          50          55          60
93 Ile Glu Phe Val Ser Ser Ala Tyr Ser Asp Arg Pro Glu Ser Phe Met
94 65          70          75          80
97 Asp Lys Arg Arg Leu Met Tyr Gly Arg Val Phe Lys Ser His Ile Phe
98          85          90          95
101 Gly Thr Ala Thr Ile Val Ser Thr Asp Ala Glu Val Asn Arg Ala Val
102          100          105          110
105 Leu Gln Ser Asp Ser Thr Ala Phe Val Pro Phe Tyr Pro Lys Thr Val
106          115          120          125
109 Arg Glu Leu Met Gly Lys Ser Ser Ile Leu Leu Ile Asn Gly Ser Leu
110          130          135          140
113 His Arg Arg Phe His Gly Leu Val Gly Ser Phe Leu Lys Ser Pro Leu
114 145          150          155          160
117 Leu Lys Ala Gln Ile Val Arg Asp Met His Lys Phe Leu Ser Glu Ser
118          165          170          175
121 Met Asp Leu Trp Ser Glu Asp Gln Pro Val Leu Leu Gln Asp Val Ser
122          180          185          190
125 Lys Thr Val Ala Phe Lys Val Leu Ala Lys Ala Leu Ile Ser Val Glu
126          195          200          205
129 Lys Gly Glu Asp Leu Glu Glu Leu Lys Arg Glu Phe Glu Asn Phe Ile
130          210          215          220
133 Ser Gly Leu Met Ser Leu Pro Ile Asn Phe Pro Gly Thr Gln Leu His
134 225          230          235          240
137 Arg Ser Leu Gln Ala Lys Lys Asn Met Val Lys Gln Val Glu Arg Ile
138          245          250          255
141 Ile Glu Gly Lys Ile Arg Lys Thr Lys Asn Lys Glu Glu Asp Asp Val
142          260          265          270
145 Ile Ala Lys Asp Val Val Asp Val Leu Leu Lys Asp Ser Ser Glu His
146          275          280          285
149 Leu Thr His Asn Leu Ile Ala Asn Asn Met Ile Asp Met Met Ile Pro
150          290          295          300
153 Gly His Asp Ser Val Pro Val Leu Ile Thr Leu Ala Val Lys Phe Leu
154 305          310          315          320
157 Ser Asp Ser Pro Ala Ala Leu Asn Leu Leu Thr Lys Asn Met Lys Leu
158          325          330          335
161 Lys Ser Leu Lys Glu Leu Thr Gly Glu Pro Leu Tyr Trp Asn Asp Tyr
162          340          345          350
165 Leu Ser Leu Pro Leu Thr Gln Lys Val Ile Thr Glu Thr Leu Arg Met
166          355          360          365
169 Gly Asn Val Ile Ile Gly Val Met Arg Lys Ala Met Lys Asp Val Glu
170          370          375          380
173 Ile Lys Gly Tyr Val Ile Pro Lys Gly Trp Cys Phe Leu Ala Tyr Leu
174 385          390          395          400

```

## RAW SEQUENCE LISTING

DATE: 09/21/2004

PATENT APPLICATION: US/10/507,106

TIME: 09:27:01

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

```

177 Arg Ser Val His Leu Asp Glu Ala Tyr Tyr Glu Ser Pro Tyr Lys Phe
178           405           410           415
181 Asn Pro Trp Arg Trp Gln Glu Arg Asp Met Asn Thr Ser Ser Phe Ser
182           420           425           430
185 Pro Phe Gly Gly Gly Gln Arg Leu Cys Pro Gly Leu Asp Leu Ala Arg
186           435           440           445
189 Leu Glu Thr Ser Val Phe Leu His His Leu Val Thr Arg Phe Arg Trp
190           450           455           460
193 Ile Ala Glu Glu Asp Thr Ile Ile Asn Phe Pro Thr Val His Met Lys
194 465           470           475           480
197 Asn Lys Leu Pro Ile Trp Ile Lys Arg Ile
198           485           490
201 <210> SEQ ID NO: 3
202 <211> LENGTH: 1934
203 <212> TYPE: DNA
204 <213> ORGANISM: Arabidopsis thaliana
206 <220> FEATURE:
207 <221> NAME/KEY: misc_feature
208 <222> LOCATION: (1748)..(1748)
209 <223> OTHER INFORMATION: n means A, C, G or T.
212 <400> SEQUENCE: 3
213 tgtgcttagg catatagtta ttcccaagaa accggttttaa ctgtttacgt atgcaacctc      60
215 cggcaagcgc aggacttttc cggtcgcccg aaaatctccc ttggccttat aattacatgg      120
217 attatttggt cgctggtttc ttggttttga cggccggaat acttctccgt ccatggctct      180
219 ggtttcgtct acgaaactcg aaaacgaaag atggagatga agaagaagat aatgaggaga      240
221 agaagaaggg aatgattcca aacggaagct taggctggcc ggtgatcgga gaaaccctaa      300
223 acttcatcgc ttgtggttat tcttctcggc ctgttacctt catggacaaa cgaaagtctt      360
225 tatacgggaa agtggtcaaa acgaacataa tagggacacc aatcataata tcaaccgatg      420
227 cagaggtgaa taaagtgggtg ctccaaaacc atgggaacac atttgtccct gcatacccta      480
229 aatcaattac ggaactactt ggagaaaact ctattctcag catcaatgga cctcatcaaa      540
231 aaaggcttca cacgctcatt ggcgcgttcc tcagatctcc tcacctcaaa gaccggatca      600
233 ctcgagacat tgaggcctcg gttgttctca ctttggcgctc ttgggctcaa cttccattgg      660
235 ttcattgttc ggatgagatc aaaaagatga cgtttgatg attagtataa gtgttgatga      720
237 gcacatctcc tggatgaagat atgaacattc tcaaacttga gttcgaagaa ttcattcaaa      780
239 gtttgatttg tatcccaatc aaattccctg gcactagact ctacaaatcc ttaaaggcga      840
241 aagagagggtt aataaagatg gtaaaaaagg ttgtggagga gagacaagtg gcatgacaa      900
243 cgacgtctcc ggcaaagtac gtggtggacg tacttctaag agacggtggt gattcagaga      960
245 agcaatctca accgtcagat ttcgtcagcg gaaagatcgt agagatgatg ataccgggag     1020
247 aggaaacaat gccaacggcg atgaccttgg ctgtcaaatt cttaagtga aaccccgctc     1080
249 ctctagccaa actcgtggag gagaatatgg agatgaagag gcgtaaattg gaattgggag     1140
251 aagaatacaa gtggaccgat tatatgtctc tctcttttac tcaaaatgtg ataaacgaaa     1200
253 cgcttagaat ggctaacatt attaacgggg tgtggaggaa agctctcaag gatgtagaaa     1260
255 ttaaagggtta cttaataccg aaaggatggg gtgtattggc atcattcata tcggttcaca     1320
257 tggatgaaga catttatgat aatccctatc aattcgatcc gtggagatgg gacagaatta     1380
259 atggatcggc aaacagcagt atttgcttca caccctttgg tgggtgggcaa aggctatgtc     1440
261 ctggtttaga gctgtcgaag ctcgaaatat ccattcttct tcaccacctt gtaaccgggt     1500
263 acagttggac ggctgaggaa gacgagatag tgtcatttcc gactgtgaag atgaagcgga     1560
265 ggctcccgat ccgagtggct actgtagatg atagtgttcc tccgatctca cttgaagatc     1620
267 attaatagat catttcaaa aacaaaactg tttgtgcaaa gaggaagcag agaagtaaac     1680

```

## RAW SEQUENCE LISTING

DATE: 09/21/2004

PATENT APPLICATION: US/10/507,106

TIME: 09:27:01

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

```

269 aaatgatctt attaacaaat agtagagaag agaagcaaac aagattgggtg ggtaagacag 1740
W--> 271 aaagaacnaa acgtacagct agtgatggct caaagatgag agattctaata tataattttt 1800
273 tttgtttgtc atgtcaaatt ataagcgttg gttaggttgt ccctttctct tttatttatc 1860
275 gtaccaaacg caagttgaga tatgattcca tatatatgga tgatagatat gtatattaat 1920
277 atatagcggc cggg 1934
280 <210> SEQ ID NO: 4
281 <211> LENGTH: 524
282 <212> TYPE: PRT
283 <213> ORGANISM: Arabidopsis thaliana
285 <400> SEQUENCE: 4
287 Met Gln Pro Pro Ala Ser Ala Gly Leu Phe Arg Ser Pro Glu Asn Leu
288 1 5 10 15
291 Pro Trp Pro Tyr Asn Tyr Met Asp Tyr Leu Val Ala Gly Phe Leu Val
292 20 25 30
295 Leu Thr Ala Gly Ile Leu Leu Arg Pro Trp Leu Trp Phe Arg Leu Arg
296 35 40 45
299 Asn Ser Lys Thr Lys Asp Gly Asp Glu Glu Glu Asp Asn Glu Glu Lys
300 50 55 60
303 Lys Lys Gly Met Ile Pro Asn Gly Ser Leu Gly Trp Pro Val Ile Gly
304 65 70 75 80
307 Glu Thr Leu Asn Phe Ile Ala Cys Gly Tyr Ser Ser Arg Pro Val Thr
308 85 90 95
311 Phe Met Asp Lys Arg Lys Ser Leu Tyr Gly Lys Val Phe Lys Thr Asn
312 100 105 110
315 Ile Ile Gly Thr Pro Ile Ile Ile Ser Thr Asp Ala Glu Val Asn Lys
316 115 120 125
319 Val Val Leu Gln Asn His Gly Asn Thr Phe Val Pro Ala Tyr Pro Lys
320 130 135 140
323 Ser Ile Thr Glu Leu Leu Gly Glu Asn Ser Ile Leu Ser Ile Asn Gly
324 145 150 155 160
327 Pro His Gln Lys Arg Leu His Thr Leu Ile Gly Ala Phe Leu Arg Ser
328 165 170 175
331 Pro His Leu Lys Asp Arg Ile Thr Arg Asp Ile Glu Ala Ser Val Val
332 180 185 190
335 Leu Thr Leu Ala Ser Trp Ala Gln Leu Pro Leu Val His Val Gln Asp
336 195 200 205
339 Glu Ile Lys Lys Met Thr Phe Glu Ile Leu Val Lys Val Leu Met Ser
340 210 215 220
343 Thr Ser Pro Gly Glu Asp Met Asn Ile Leu Lys Leu Glu Phe Glu Glu
344 225 230 235 240
347 Phe Ile Lys Gly Leu Ile Cys Ile Pro Ile Lys Phe Pro Gly Thr Arg
348 245 250 255
351 Leu Tyr Lys Ser Leu Lys Ala Lys Glu Arg Leu Ile Lys Met Val Lys
352 260 265 270
355 Lys Val Val Glu Glu Arg Gln Val Ala Met Thr Thr Thr Ser Pro Ala
356 275 280 285
359 Asn Asp Val Val Asp Val Leu Arg Asp Gly Gly Asp Ser Glu Lys
360 290 295 300
363 Gln Ser Gln Pro Ser Asp Phe Val Ser Gly Lys Ile Val Glu Met Met

```

## RAW SEQUENCE LISTING

DATE: 09/21/2004

PATENT APPLICATION: US/10/507,106

TIME: 09:27:01

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

```

364 305          310          315          320
367 Ile Pro Gly Glu Glu Thr Met Pro Thr Ala Met Thr Leu Ala Val Lys
368          325          330          335
371 Phe Leu Ser Asp Asn Pro Val Ala Leu Ala Lys Leu Val Glu Glu Asn
372          340          345          350
375 Met Glu Met Lys Arg Arg Lys Leu Glu Leu Gly Glu Glu Tyr Lys Trp
376          355          360          365
379 Thr Asp Tyr Met Ser Leu Ser Phe Thr Gln Asn Val Ile Asn Glu Thr
380          370          375          380
383 Leu Arg Met Ala Asn Ile Ile Asn Gly Val Trp Arg Lys Ala Leu Lys
384 385          390          395          400
387 Asp Val Glu Ile Lys Gly Tyr Leu Ile Pro Lys Gly Trp Cys Val Leu
388          405          410          415
391 Ala Ser Phe Ile Ser Val His Met Asp Glu Asp Ile Tyr Asp Asn Pro
392          420          425          430
395 Tyr Gln Phe Asp Pro Trp Arg Trp Asp Arg Ile Asn Gly Ser Ala Asn
396          435          440          445
399 Ser Ser Ile Cys Phe Thr Pro Phe Gly Gly Gly Gln Arg Leu Cys Pro
400          450          455          460
403 Gly Leu Glu Leu Ser Lys Leu Glu Ile Ser Ile Phe Leu His His Leu
404 465          470          475          480
407 Val Thr Arg Tyr Ser Trp Thr Ala Glu Glu Asp Glu Ile Val Ser Phe
408          485          490          495
411 Pro Thr Val Lys Met Lys Arg Arg Leu Pro Ile Arg Val Ala Thr Val
412          500          505          510
415 Asp Asp Ser Ala Ser Pro Ile Ser Leu Glu Asp His
416          515          520

```

419 &lt;210&gt; SEQ ID NO: 5

420 &lt;211&gt; LENGTH: 20

421 &lt;212&gt; TYPE: DNA

422 &lt;213&gt; ORGANISM: Artificial sequence

424 &lt;220&gt; FEATURE:

425 &lt;223&gt; OTHER INFORMATION: PCR amplification primer

427 &lt;400&gt; SEQUENCE: 5

428 gttaaaacac taatggacac

20

431 &lt;210&gt; SEQ ID NO: 6

432 &lt;211&gt; LENGTH: 21

433 &lt;212&gt; TYPE: DNA

434 &lt;213&gt; ORGANISM: Artificial sequence

436 &lt;220&gt; FEATURE:

437 &lt;223&gt; OTHER INFORMATION: PCR amplification primer

439 &lt;400&gt; SEQUENCE: 6

440 tgatttatat tcttttgatc c

21

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 09/21/2004

PATENT APPLICATION: US/10/507,106

TIME: 09:27:02

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 1748 ✓

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/507,106

DATE: 09/21/2004

TIME: 09:27:02

Input Set : A:\26352U sequence listing.ST25.txt

Output Set: N:\CRF4\09212004\J507106.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:1740